

Pengantar Menuju Analisis Data Longitudinal *‘Sebenarnya’*



Wahyu Widhiarso | Fakultas Psikologi UGM

Pengantar Menuju Analisis Data Longitudinal '*Sebenarnya*'



via ~~Easy Way~~ Multilevel Modeling *)

*) dengan segenap ralat

Terminologi

Data Longitudinal

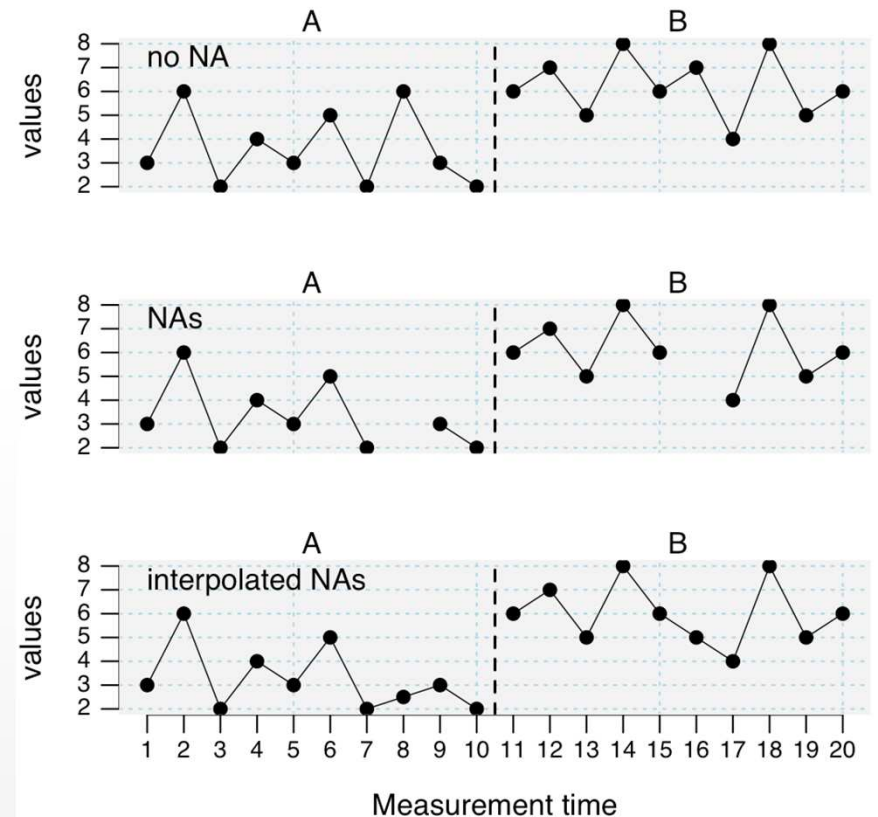
- Data yang diambil melalui beberapa pengamatan berulang (antar waktu) pada suatu unit
 - Unit dapat berupa individu, sekolah, benda dsb.
- Data longitudinal seringkali dikontraskan dengan Cross sectional data

Repeated Measures

- Biasanya dipakai dalam penelitian eksperimental dimana subjek mendapatkan perlakuan yang berbeda-beda pada suatu interval waktu
 - Misalnya desain ABBA

Kelebihan Data Longitudinal

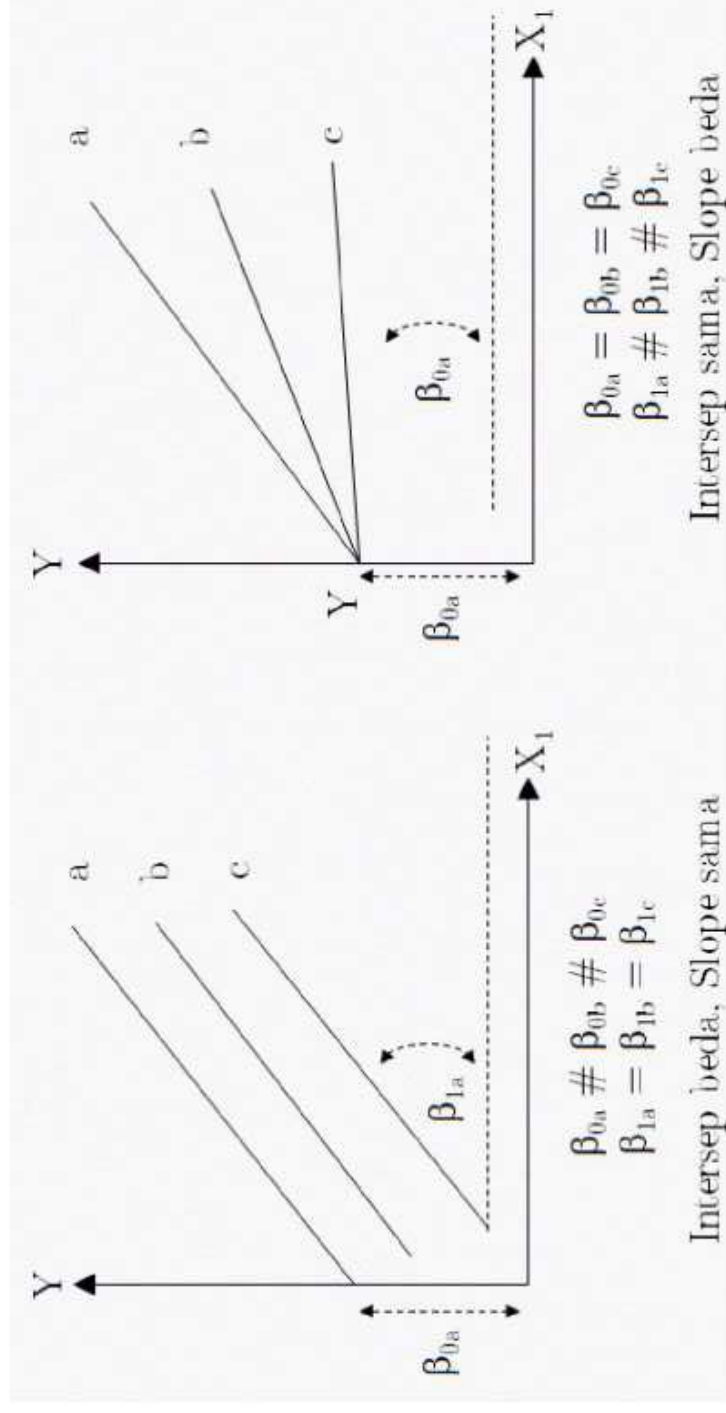
- Memberikan informasi mengenai pola perubahan
- Lebih ekonomis, tidak melibatkan banyak partisipan
- Mengurangi varians eror terkait dengan perbedaan karakteristik individu



Fixed vs Random Effect

4 Intersep dan Slope

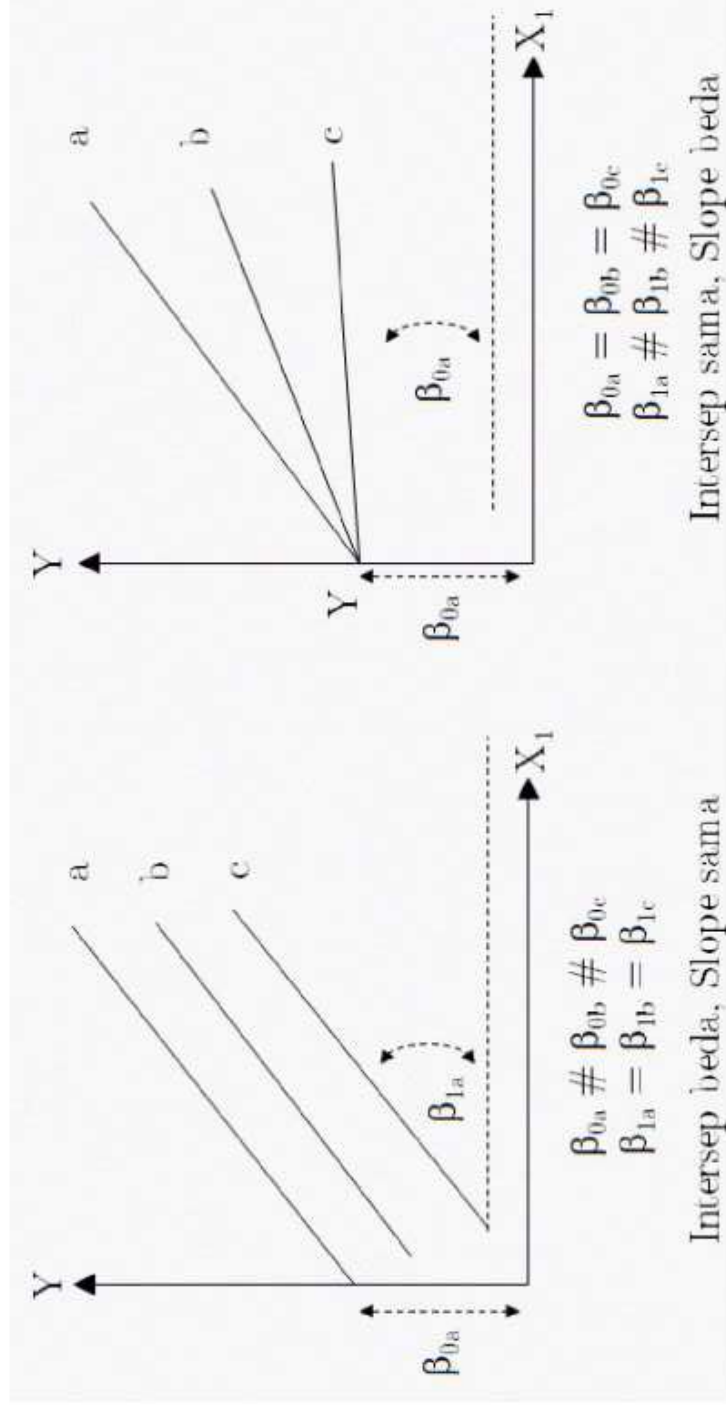
Kita telah tahu bahwa intersep adalah nilai Y ketika nilai $X=0$ yang ditunjukkan dengan perpotongan garis regresi dan sumbu Y . Slope menunjukkan kemiringan garis regresi. Gambar 4 menunjukkan garis regresi pada grup atau kelompok individu.



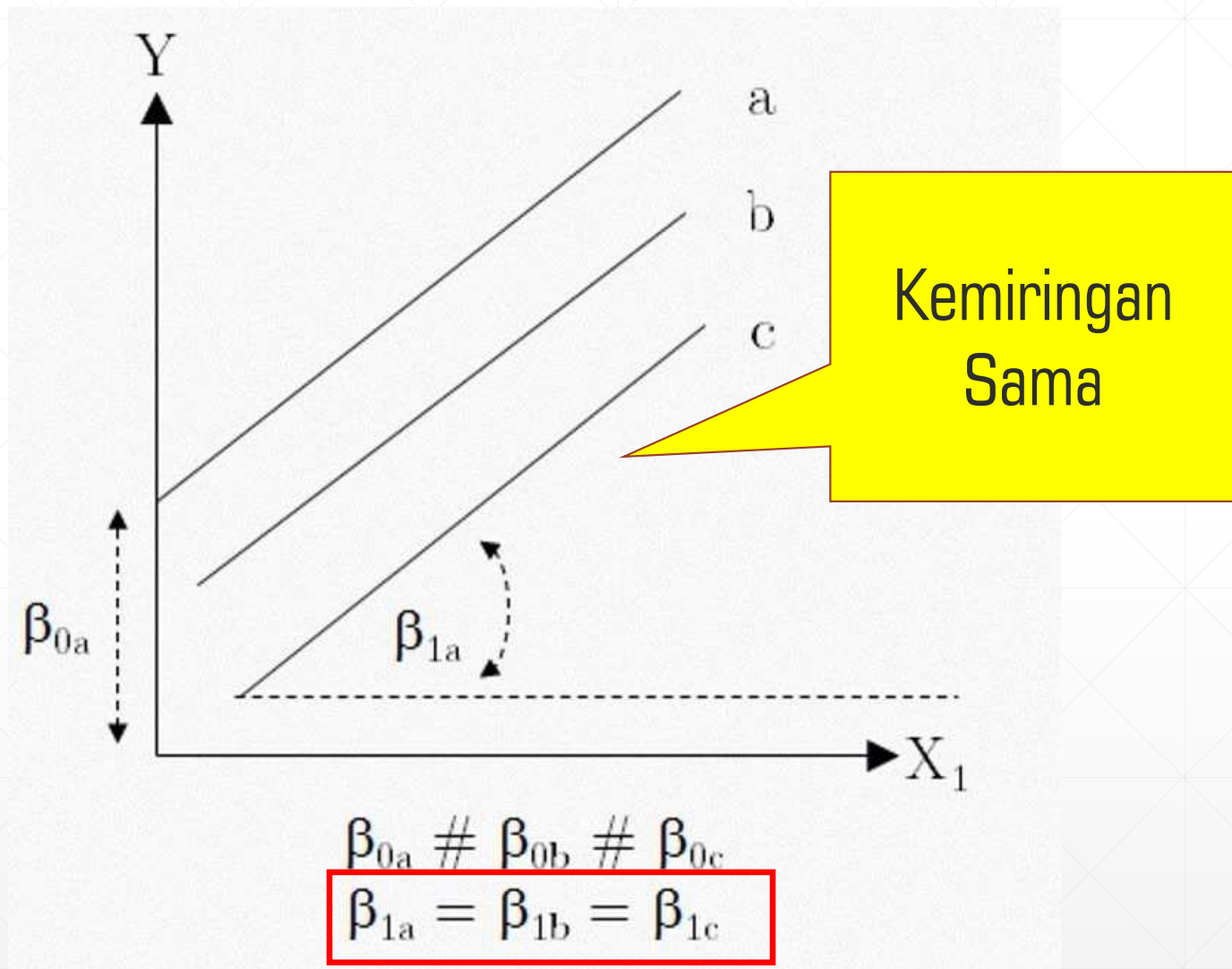
Gambar 4 Model-Model Regresi Multilevel

4 Intersep dan Slope

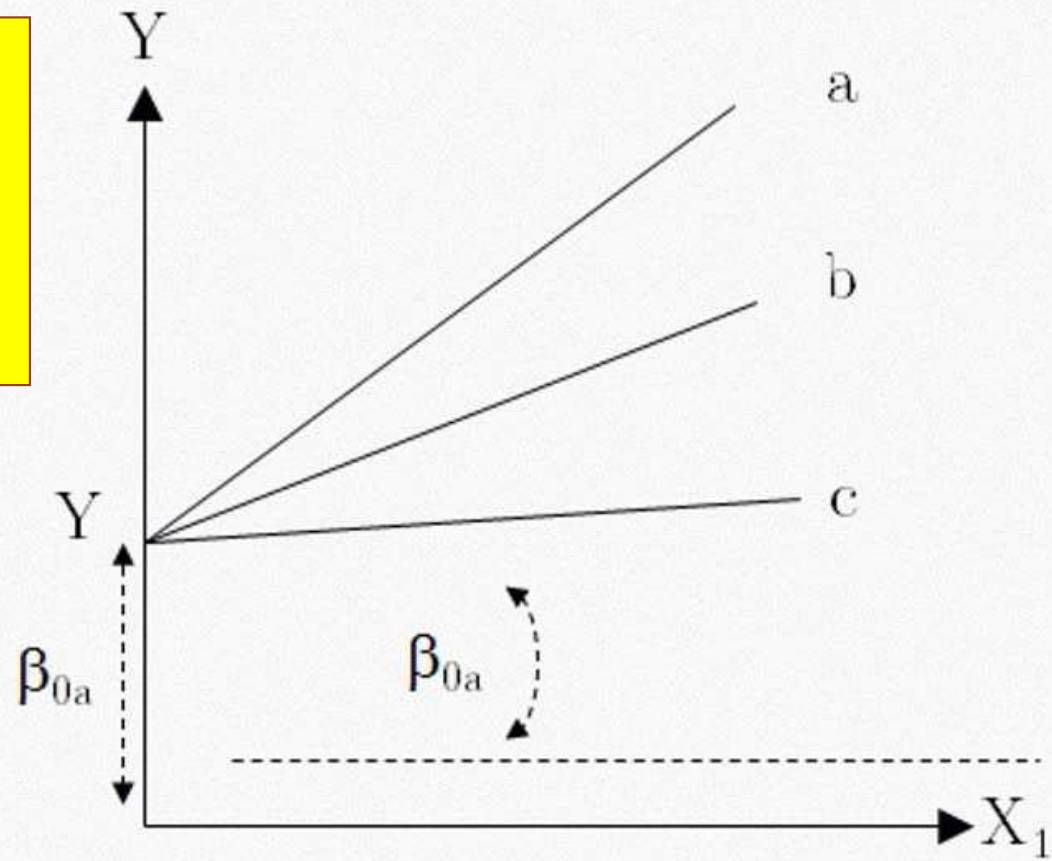
Kita telah tahu bahwa intersep adalah nilai Y ketika nilai $X=0$ yang ditunjukkan dengan perpotongan garis regresi dan sumbu Y . Slope menunjukkan kemiringan garis regresi. Gambar 4 menunjukkan garis regresi pada grup atau kelompok individu.



Gambar 4 Model-Model Regresi Multilevel

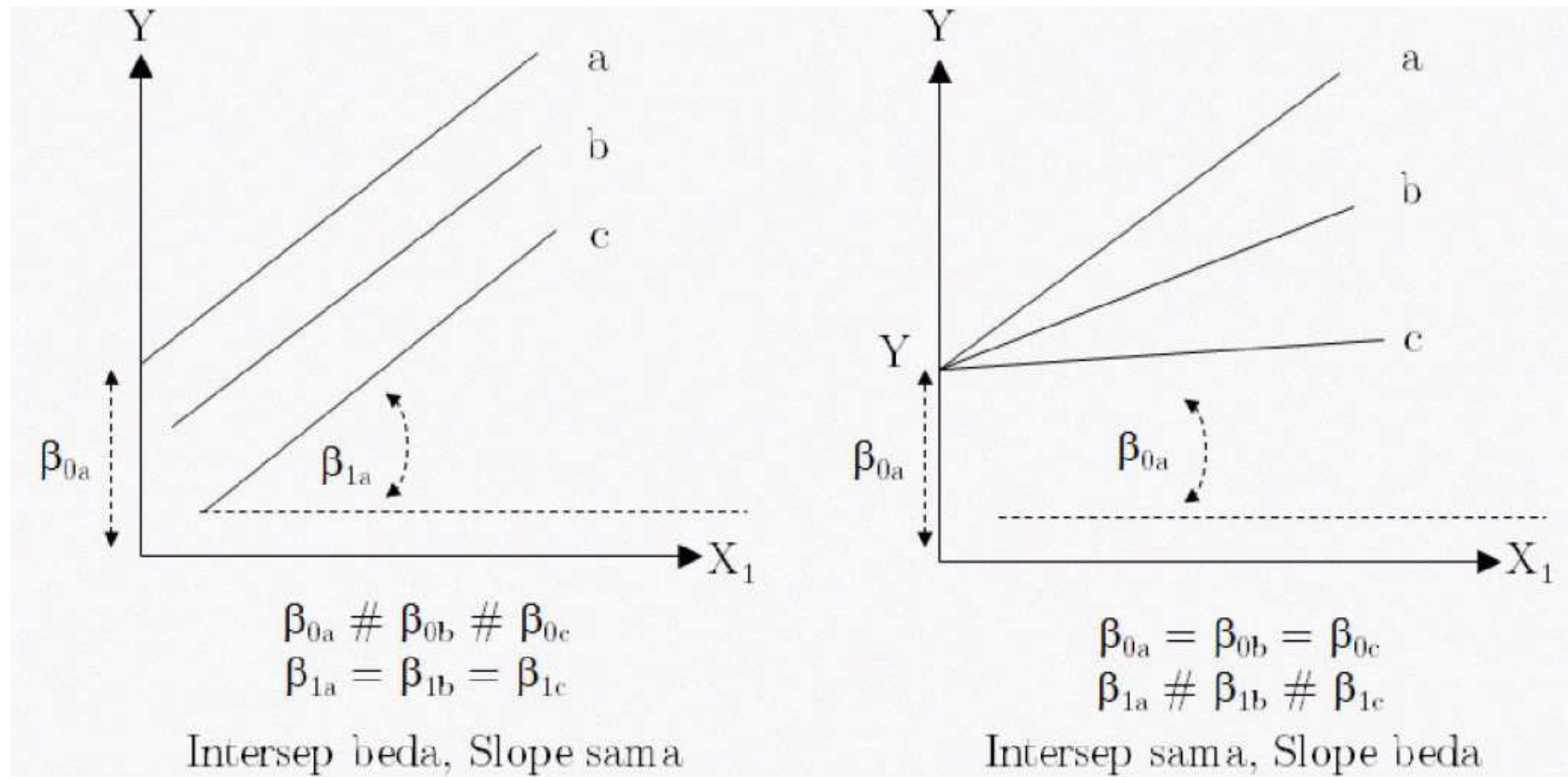


Intercept
Sama



$$\beta_{0a} = \beta_{0b} = \beta_{0c}$$

$$\beta_{1a} \neq \beta_{1b} \neq \beta_{1c}$$



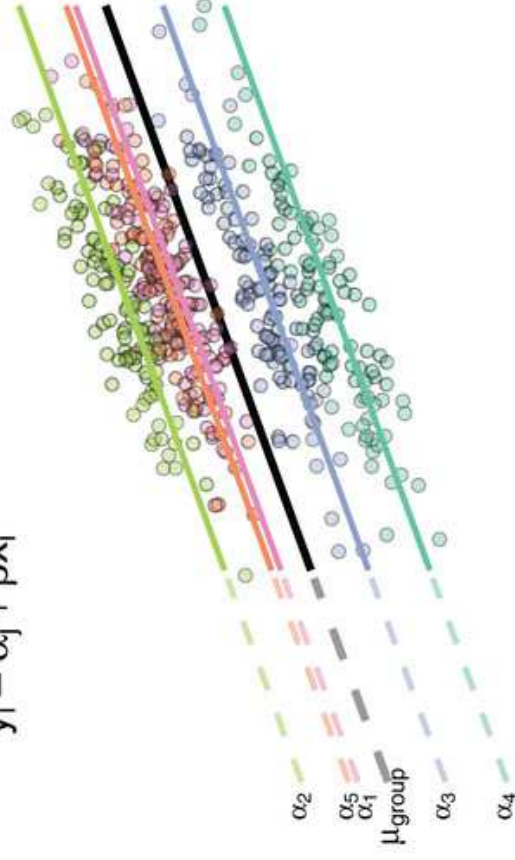
Gambar 4 Model-Model Regresi Multilevel

Model mana yang paling tepat?

A

Random Intercepts

$$y_i = \alpha_j + \beta x_i$$

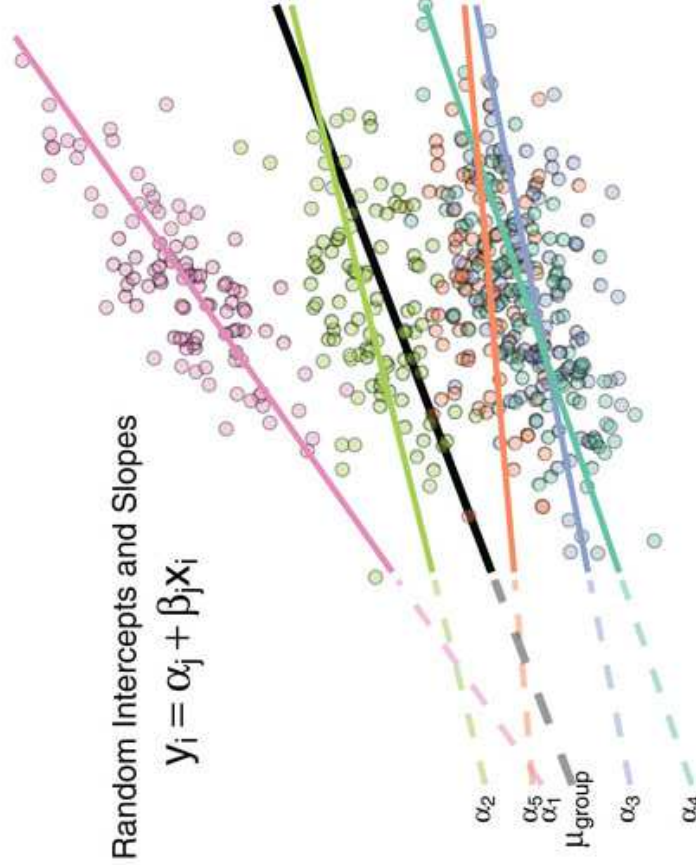


Predictor Variable x

B

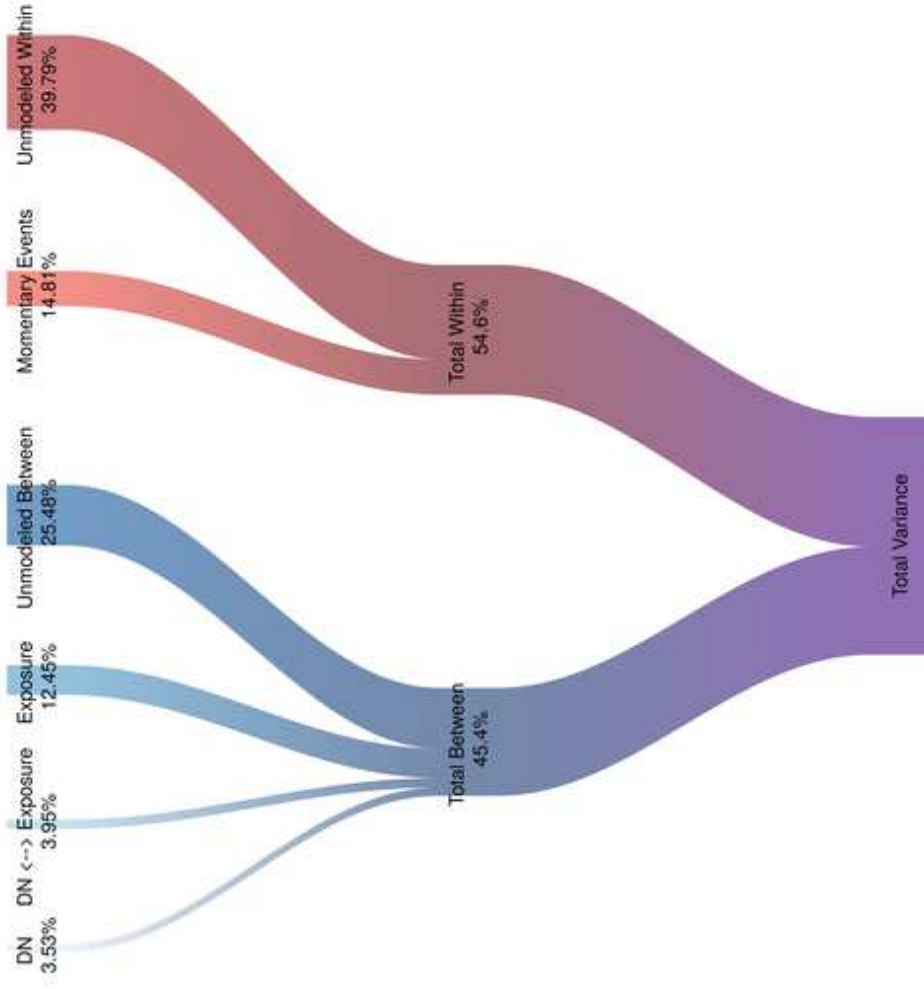
Random Intercepts and Slopes

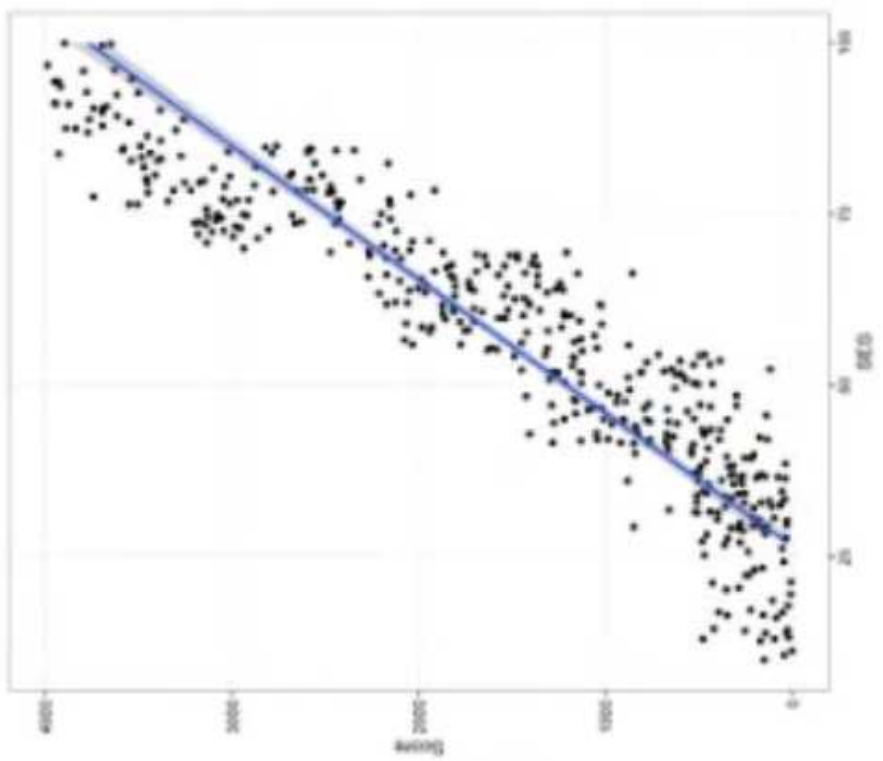
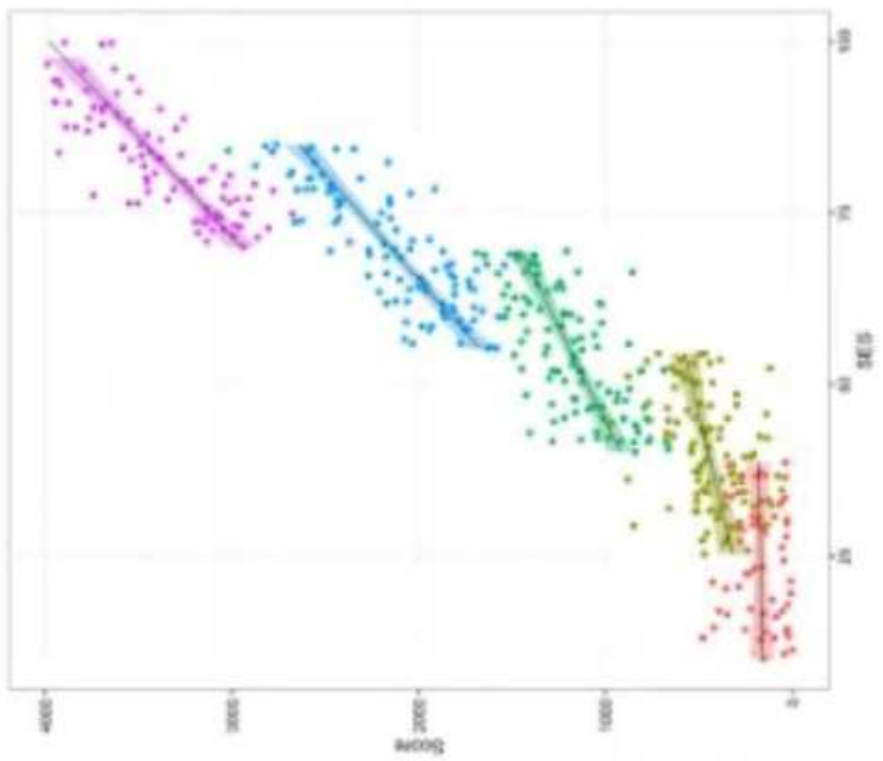
$$y_i = \alpha_j + \beta_j x_i$$



Predictor Variable x

Study 1 - Riverplot of Total Variance Decomposition Estimated from Positive Mood Models





Ketika kita terapkan analisis regresi pada kelompok kita akan mendapatkan sejumlah garis regresi yang jumlahnya sama dengan jumlah kelompok. Parameter intersep dan kelompok tersebut bisa saja berbeda.

Kedua model ini kita terapkan pada data kita.

Model mana yang paling tepat akan terlihat dari uji kecocokan model yang paling tinggi.

Persamaan regresi untuk masing-masing model adalah sebagai berikut:

Intersep sama & Slope beda

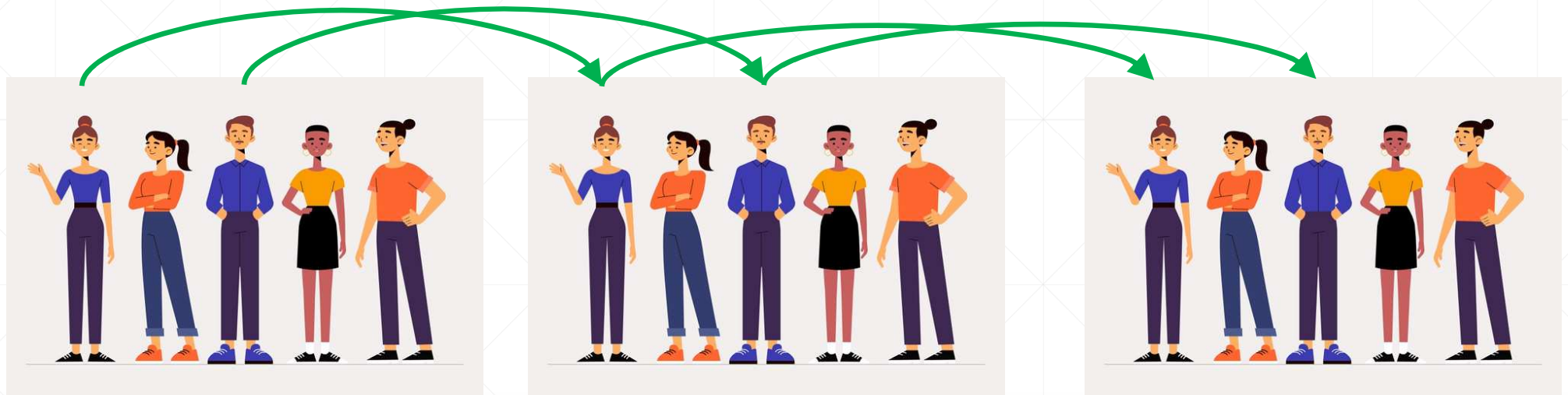
$$Y_{ij} = \beta_0 + \beta_{1j}X_{ij} + e_{ij}$$

Intersep beda & Slope sama

$$Y_{ij} = \beta_{0j} + \beta_1X_{ij} + e_{ij}$$

Repeated Measures Data

Varians Dalam Amatan



Amatan 1

Amatan 2

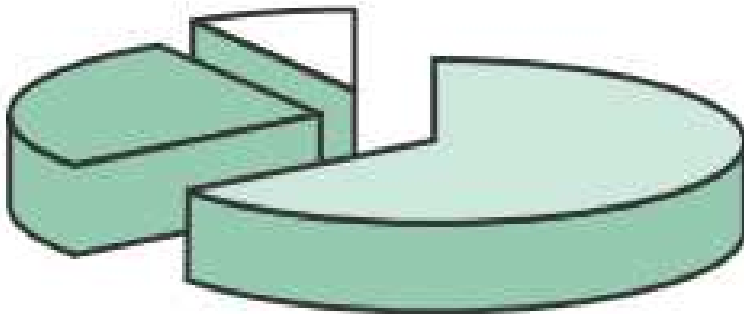
Amatan 3

Varians Antar Amatan

$\text{Varians Dalam Amatan} - \text{Varians Antar Amatan} = \text{Varians Error}$

Varians Dalam Amatan

Diharapkan sekecil-kecilnya karena menunjukkan perbedaan individual karena dipengaruhi karakteristik unik tiap individu bukan karena kelompok



- Between-groups variance
- Individual differences variance
- Residual variance

Varians Antar Amatan



Diharapkan sebesar-besarnya untuk menunjukkan perbedaan antar kelompok/amatan

Contoh Data

Subjek	T1	T2	T3	T4
A1	62	58	63	64
A2	63	60	68	65
A3	65	61	72	65
A4	68	64	58	61
A5	69	65	54	59
A6	71	67	65	50
A7	78	66	67	50
A8	75	73	75	45

Tahap 1 = Variasi Total

Subjek	T1	T2	T3	T4
A1	62	58	63	64
A2	63	60	68	65
A3	65	61	72	65
A4	68	64	58	61
A5	69	65	54	59
A6	71	67	65	50
A7	78	66	67	50
A8	75	73	75	45

**Rerata Total =
63.96**

$$JK-TOT = \sum (X_i - M_{total})$$

$$JK-TOT = (62 - 63.96)^2 + (63 - 63.96)^2 + \dots = 1705,868$$



VARIASI TOTAL

1705,868

Tahap 2 = Variasi Dalam Orang

Subjek	T1	T2	T3	T4	Rerata
A1	62	58	63	64	61.75
A2	63	60	68	65	64
A3	65	61	72	65	65.75
A4	68	64	58	61	62.75
A5	69	65	54	59	61.75
A6	71	67	65	50	63.25
A7	78	66	67	50	65.25
A8	75	73	75	45	67

Rerata A1 = 61.75

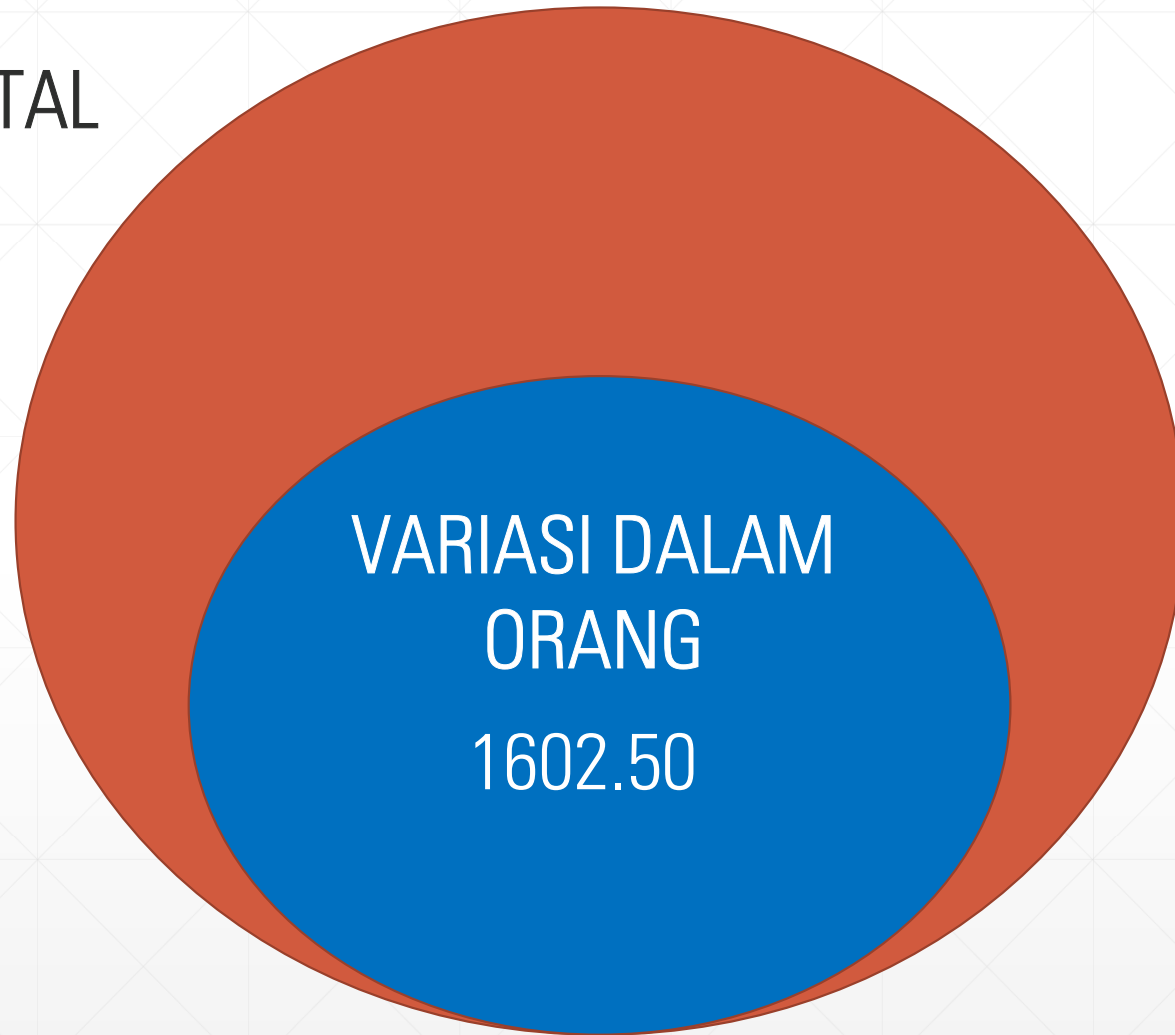
JK-DALAM ORANG 1 =

$$\Sigma(X_i - M_{\text{orang-i}})^2 =$$
$$(62-61.75)^2 + (58-61.75)^2 + \dots$$

UNTUK SETIAP ORANG

JK-DALAM ORANG = 1602.50

VARIASI TOTAL
1705,868



Tahap 3 = Variasi Antar Amatan

Subjek	T1	T2	T3	T4
A1	62	58	63	64
A2	63	60	68	65
A3	65	61	72	65
A4	68	64	58	61
A5	69	65	54	59
A6	71	67	65	50
A7	78	66	67	50
A8	75	73	75	45
	68.88	64.25	65.25	57.38

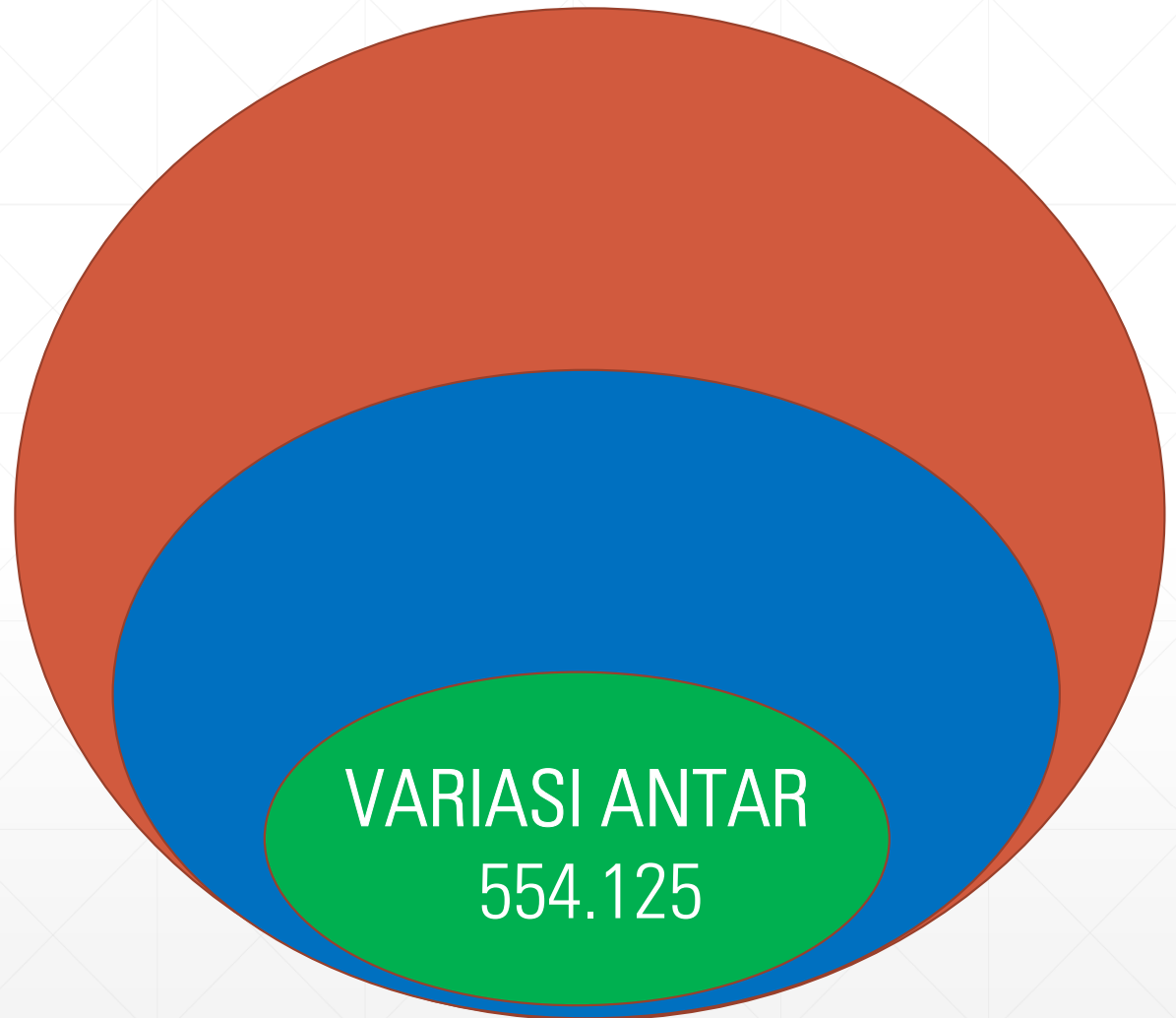
Rerata Total = 63.96

$$\begin{aligned} \text{JK-ANTAR} &= \sum (M_k - M_{\text{tot}})^2 = \\ &= (68.88 - 63.96)^2 + (58 - 63.96)^2 \\ &\quad + \dots = 554.125 \end{aligned}$$

JK-ANTAR AMATAN = 554.125

VARIASI TOTAL
1705,868

VARIASI DALAM
1602.50

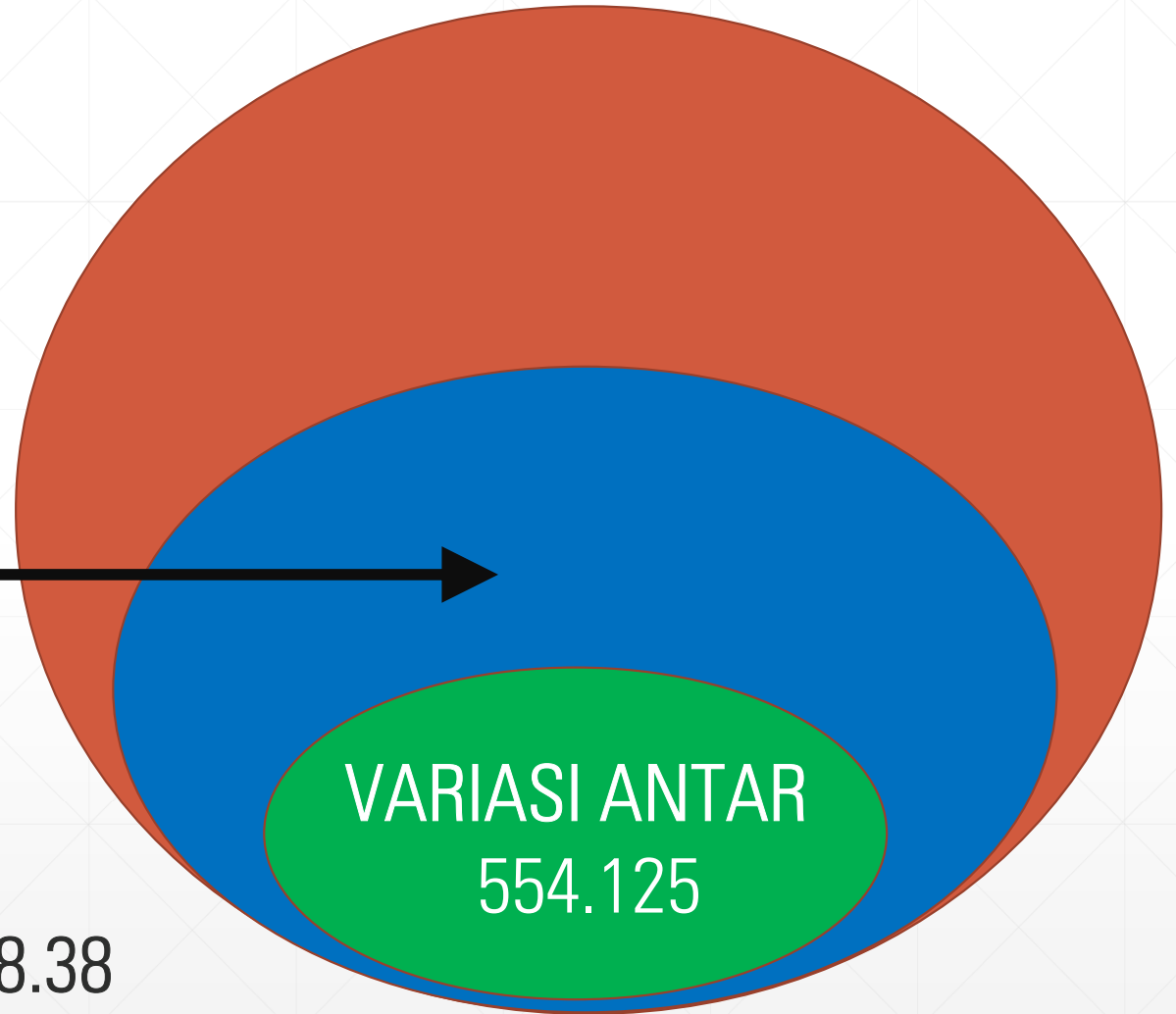


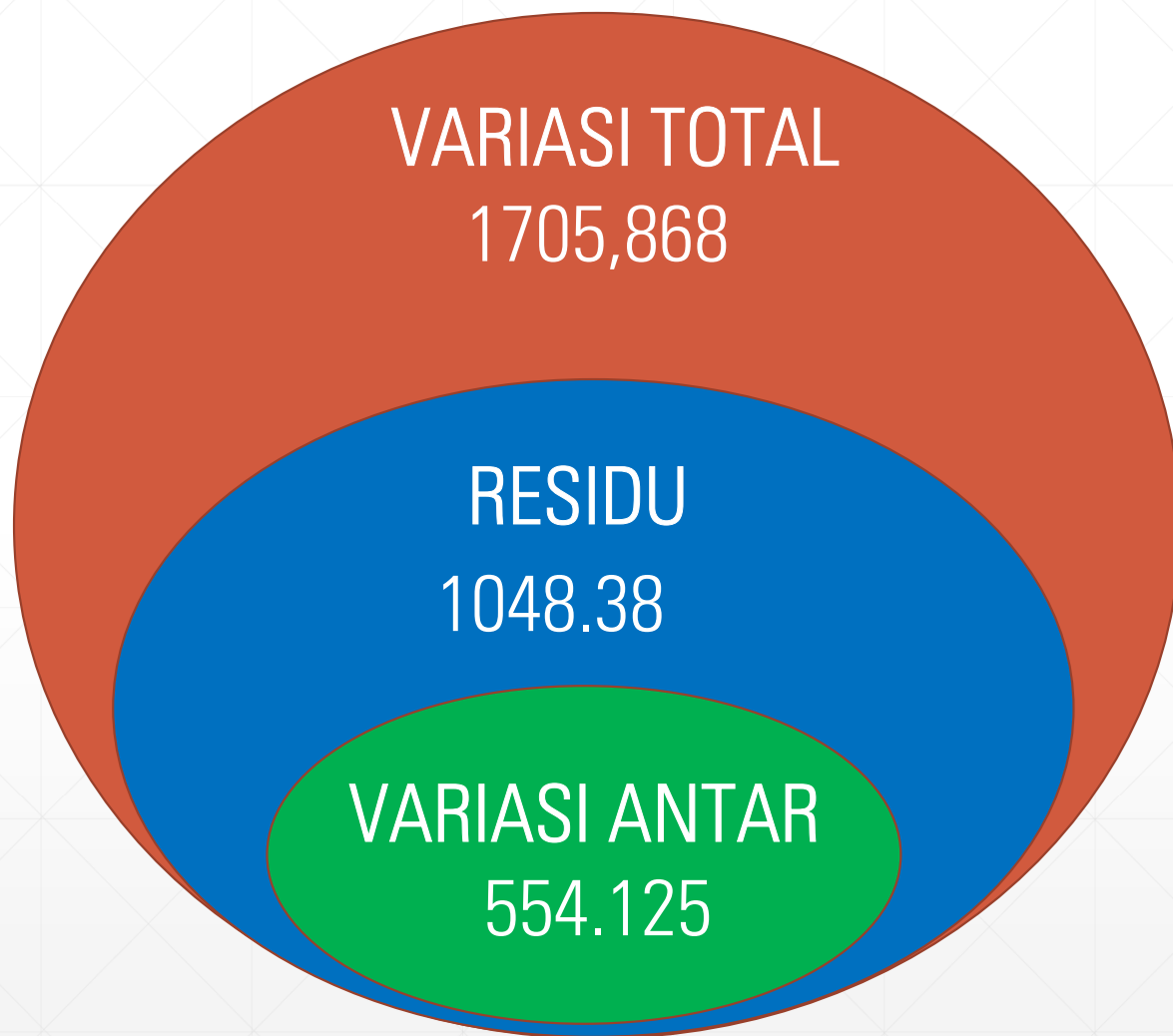
VARIASI TOTAL
1705,868

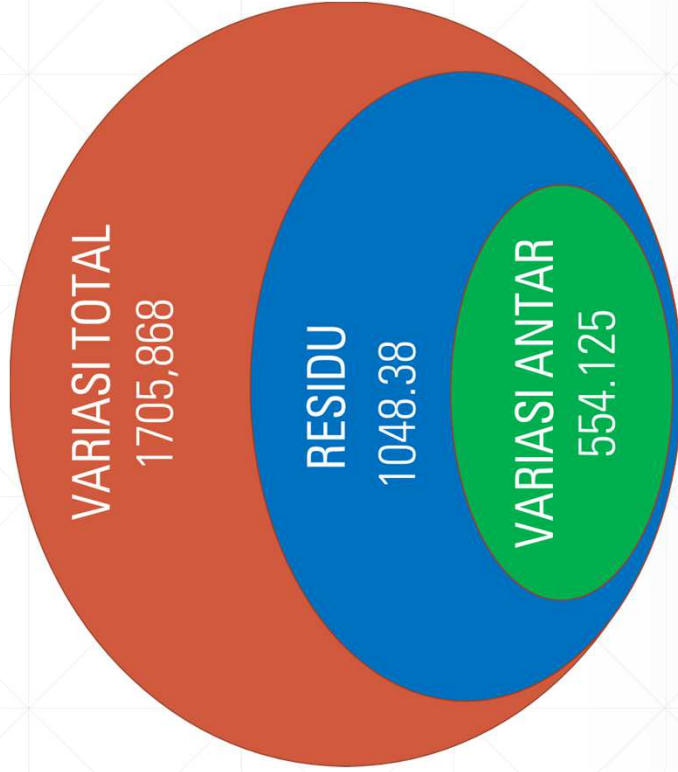
VARIASI DALAM
1602.50

EROR
(RESIDU)

$$1602.50 - 554.125 = 1048.38$$







Tests of Within-Subjects Effects

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square
Amatan			
Sphericity Assumed	554.125	3	184.708
Greenhouse-Geisser	554.125	1.673	331.245
Huynh-Feldt	554.125	2.137	259.329
Lower-bound	554.125	1.000	554.125
Error(Amatan)			
Sphericity Assumed	1048.375	21	49.923
Greenhouse-Geisser	1048.375	11.710	89.528
Huynh-Feldt	1048.375	14.957	70.091
Lower-bound	1048.375	7.000	149.768

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Amatan	Sphericity Assumed	554.125	3	184.708	3.700	.028
	Greenhouse-Geisser	554.125	1.673	331.245	3.700	.063
	Huynh-Feldt	554.125	2.137	259.329	3.700	.047
	Lower-bound	554.125	1.000	554.125	3.700	.096
Error(Amatan)	Sphericity Assumed	1048.375	21	49.923		
	Greenhouse-Geisser	1048.375	11.710	89.528		
	Huynh-Feldt	1048.375	14.957	70.091		
	Lower-bound	1048.375	7.000	149.768		

Antara T1, T2, T3 dan T4 didapatkan
Perbedaan yang signifikan
($F = 3.70$; $p < 0.05$)

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Amatan	Sphericity Assumed	554.125	3	184.708	3.700	.028
	Greenhouse-Geisser	554.125	1.673	331.245	3.700	.063
	Huynh-Feldt	554.125	2.137	259.329	3.700	.047
	Lower-bound	554.125	1.000	554.125	3.700	.096
Error(Amatan)	Sphericity Assumed	1048.375	21	49.923		
	Greenhouse-Geisser	1048.375	11.710	89.528		
	Huynh-Feldt	1048.375	14.957	70.091		
	Lower-bound	1048.375	7.000	149.768		

Mean Square = Sum of Square / df

F = Mean Square Amatan / Mean Square Error

Tabel Hasil Analisis Anava Amatan Ulang Secara Lengkap

Table 3. Summary table for the one-way repeated measures ANOVA

	Sum of squares	Degree of freedom	Mean squares	F	Effect size (partial eta-squared)
Subject	704.96	$n - 1 = 7$	100.78		
Conditions	152.33 (SSB)	$K - 1 = 2$	76.16	15.31 ($p < 0.0001$)	$152.33 / (152.33 + 69.67) = 0.69$
Residual	69.67 (SSE)	$(n - 1)(k - 1) = 14$	4.98	(MS by condition / MS Residual)	[SS by condition / (SS by condition + SS residual)]
Total	926.96 (SST)	$(n)(k) - 1 = 24 - 1 = 23$			

Longitudinal Data

Format Data

	Subject	Gender	Region	Education	LDLF	LDHF	HDLF	HDHF
1	1	Female	North America	some	6.0	6.0	9.0	10.0
2	2	Female	North America	advance	10.0		10.0	10.0
3	3	Female	Europe	college	5.0	10.0	10.0	10.0
4	4	Female	North America	college	6.0	9.0	6.0	9.0
5	5	Female	North America	some	3.0	6.5	5.5	8.5
6	6	Female	Europe	some	2.0	0.5	7.5	3.0
7	7	Female	North America	some	10.0	10.0	10.0	10.0
8	8	Female	North America	high	10.0	10.0	9.0	10.0
9	9	Female	North America	high	9.5	9.5	6.0	10.0
10	10	Female	Other	high	8.5	10.0	7.0	
11	11	Male	North America	some	0.0	2.5	0.0	0.0
12	12	Male	North America	some	9.5	7.5	8.5	10.0
13	13	Female	North America	less	7.0	8.5	6.5	8.5
14	14	Female	North America	some	3.0	8.0	4.0	8.5
15	15	Male	North America	some	4.0	5.0	6.0	7.0
16	16	Male	North America	high	7.0	8.5	8.5	9.0
17	17	Male	North America	some	2.0	6.0	10.0	10.0
18	18	Female	North America	high	7.5	6.0	7.5	10.0

Width Format

	Subject	Gender	Region	Education	Disgust	Fright	Rating
1	1	Female	North America	some	Low	Low	6.0
2	1	Female	North America	some	Low	High	6.0
3	1	Female	North America	some	High	Low	9.0
4	1	Female	North America	some	High	High	10.0
5	2	Female	North America	advance	Low	Low	10.0
6	2	Female	North America	advance	Low	High	
7	2	Female	North America	advance	High	Low	10.0
8	2	Female	North America	advance	High	High	10.0
9	3	Female	Europe	college	Low	Low	5.0
10	3	Female	Europe	college	Low	High	10.0
11	3	Female	Europe	college	High	Low	10.0
12	3	Female	Europe	college	High	High	10.0
13	4	Female	North America	college	Low	Low	6.0
14	4	Female	North America	college	Low	High	9.0
15	4	Female	North America	college	High	Low	6.0
16	4	Female	North America	college	High	High	9.0

Long Format

Format Data

Width Format

Long Format

Data						
Analyses						
Edit						
Clipboard						
Edit						
Variables						
Rows						
id	grup	Amatan	depresi	putusasa		
1	1	Perlakuan	Pretest	23	34	
2	1	Perlakuan	Posttest	21	33	
3	1	Perlakuan	Followup	21	19	
4	2	Perlakuan	Pretest	23	37	
5	2	Perlakuan	Posttest	12	31	
6	2	Perlakuan	Followup	17	20	
7	3	Perlakuan	Pretest	24	31	
8	3	Perlakuan	Posttest	18	33	
9	3	Perlakuan	Followup	20	18	
10	4	Perlakuan	Pretest	28	30	
11	4	Perlakuan	Posttest	25	31	
12	4	Perlakuan	Followup	20	18	
13	5	Perlakuan	Pretest	25	24	
14	5	Perlakuan	Posttest	17	29	
15	5	Perlakuan	Followup	20	17	
16	6	Perlakuan	Pretest	24	42	
17	6	Perlakuan	Posttest	21	43	
18	6	Perlakuan	Followup	16	18	
19	7	Perlakuan	Pretest	25	34	
20	7	Perlakuan	Posttest	18	29	

Data

Analyses

Edit

Paste

Clipboard

Edit

Setup

Compute

Transform

Variables

Add

Delete

Filters

Add

Delete

Rows

	id	dep1	dep2	dep3	put1	put2	put3	grup	
1	1	23	21	21	34	33	19	Perlakuan	
2	2	23	12	17	37	31	20	Perlakuan	
3	3	24	18	20	31	33	18	Perlakuan	
4	4	28	25	20	30	31	18	Perlakuan	
5	5	25	17	20	24	29	17	Perlakuan	
6	6	24	21	16	42	43	18	Perlakuan	
7	7	25	18	16	34	29	18	Perlakuan	
8	8	26	16	17	43	43	18	Perlakuan	
9	9	23	15	22	38	33	20	Perlakuan	
10	10	27	18	19	40	31	21	Perlakuan	
11	11	29	16	23	29	26	23	Perlakuan	
12	12	25	20	21	35	27	20	Perlakuan	
13	13	28	20	21	26	27	20	Perlakuan	
14	14	24	19	19	37	29	18	Perlakuan	
15	15	23	25	20	36	30	20	Perlakuan	
16	16	26	20	21	37	32	18	Perlakuan	
17	17	24	19	19	32	28	20	Perlakuan	
18	18	25	18	22	32	30	22	Perlakuan	
19	19	24	20	21	27	28	26	Perlakuan	
20	20	23	16	18	36	29	20	Perlakuan	
21	21	27	18	20	37	31	23	Perlakuan	
22	22	23	17	20	38	35	22	Perlakuan	
23	23	29	18	22	34	32	28	Perlakuan	
24	24	29	22	20	30	25	26	Perlakuan	
25	25	24	16	22	38	25	25	Perlakuan	
26	26	23	26	29	34	34	24	Kontrol	
27	27	27	30	27	35	37	25	Kontrol	

Mixed Model

Dependent Variable:

Factors:

Covariates:

Cluster variables:

Estimation: ☒ REML ☒ Confidence intervals Interval: %



Fixed Effects

Components: grup Amatan putusasa

Model Terms: grup Amatan putusasa grup * Amatan

☒ Fixed Intercept

Random Effects

Components: grup | id Amatan | id putusasa | id grup : Amatan | id grup : putusasa | id Amatan : putusasa | id grup : Amatan : putusasa | id

Random Coefficients: Intercept | id

Effects correlation

Tests

Fixed Effect

Fixed Effect Omnibus tests

	F	Num df	Den df	p
grup	129.23	1	55.9	< .001
Amatan	26.00	2	107.9	< .001
putusasa	1.52	1	138.7	0.220
grup * Amatan	58.35	2	96.3	< .001

Note. Satterthwaite method for degrees of freedom

Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	23.5067	0.1942	23.126	23.8873	46.3	121.03	< .001
grup1	Perlakuan - Kontrol	-4.7115	0.4145	-5.524	-3.8992	55.9	-11.37	< .001
Amatan1	Posttest - Pretest	-3.0073	0.4174	-3.825	-2.1892	95.9	-7.20	< .001
Amatan2	Followup - Pretest	-1.9717	0.6178	-3.183	-0.7609	121.9	-3.19	0.002
putusasa	putusasa	-0.0548	0.0444	-0.142	0.0323	138.7	-1.23	0.220
grup1 * Amatan1	Perlakuan - Kontrol * Posttest - Pretest	-7.6465	0.8275	-9.268	-6.0246	95.1	-9.24	< .001
grup1 * Amatan2	Perlakuan - Kontrol * Followup - Pretest	-8.2617	0.8622	-9.952	-6.5718	98.9	-9.58	< .001

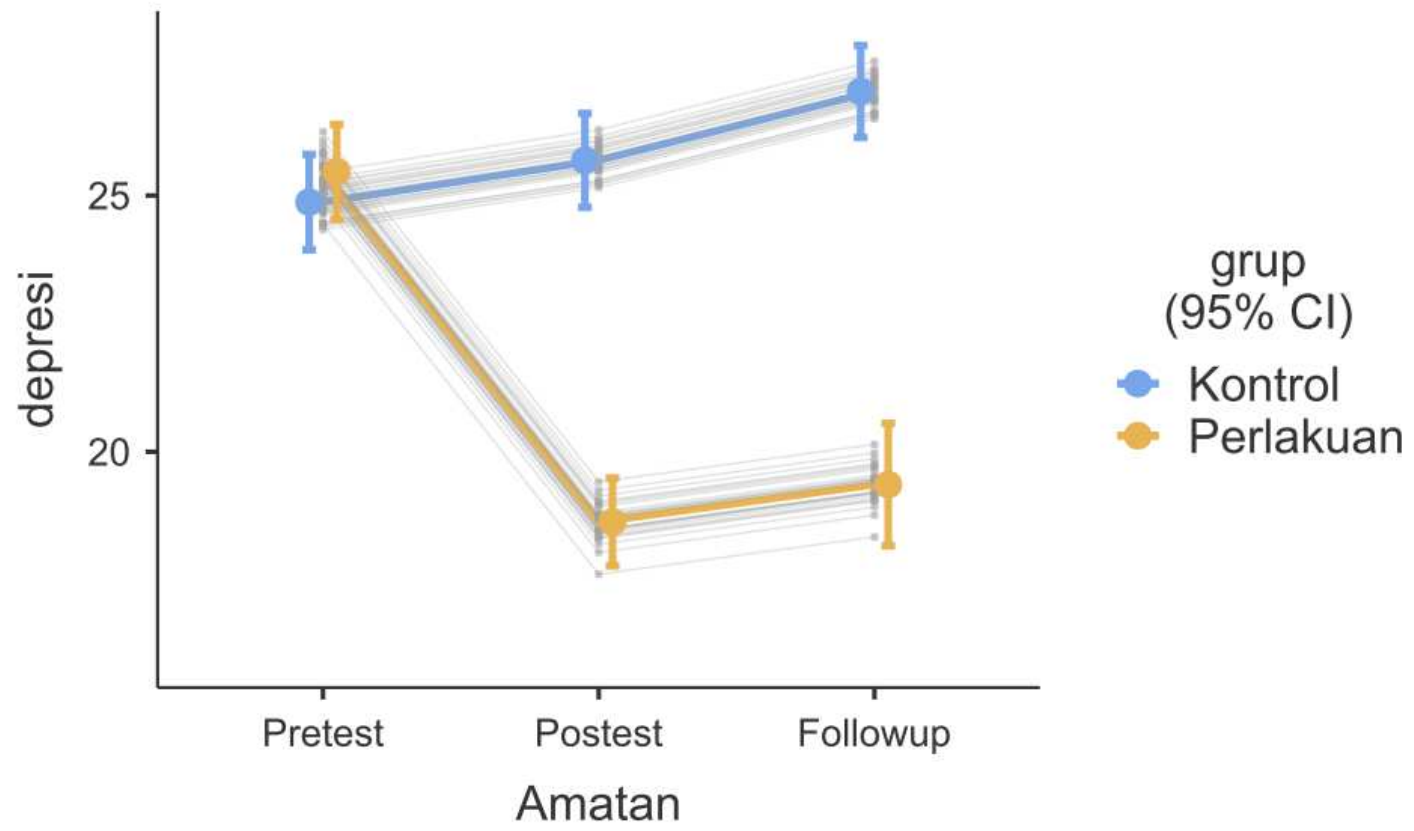
Random Effect

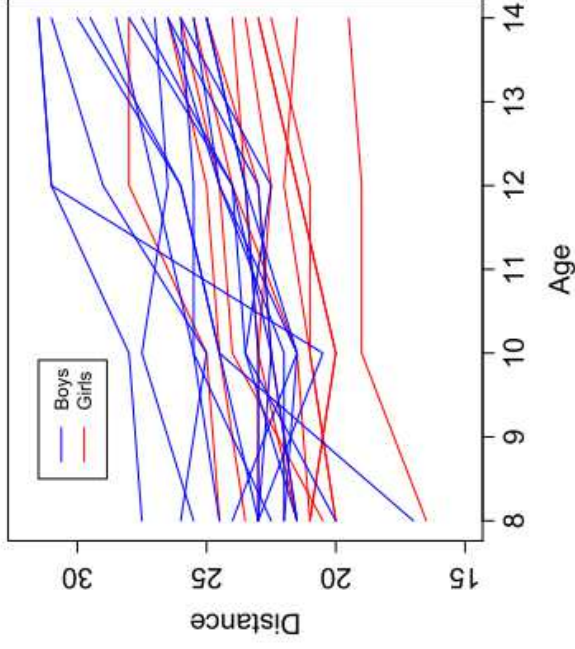
Random Components

Groups	Name	SD	Variance	ICC
id	(Intercept)	0.705	0.497	0.107
Residual		2.041	4.166	

Note. Number of Obs: 150 , groups: id 50

Plot





- Linear model plausible, but height differences between girls and boys. Growth at same speed??

- Model with interaction:

$$E(Y_{ij}) = \beta_0 + \beta_1 t_{ij} + \beta_2 G_i + \beta_3 G_i t_{ij},$$

$\mathbf{X}_i$ contains rows $[1 \ G_i \ t_{ij} \ G_i t_{ij}]$.

The data for 100 patients, 48 in the TAU group and 52 in the BtB group, are shown in [Table 8.1](#). These data are a subset of the original and are used with the kind permission of the organizers of the study, in particular Dr. Judy Proudfoot.

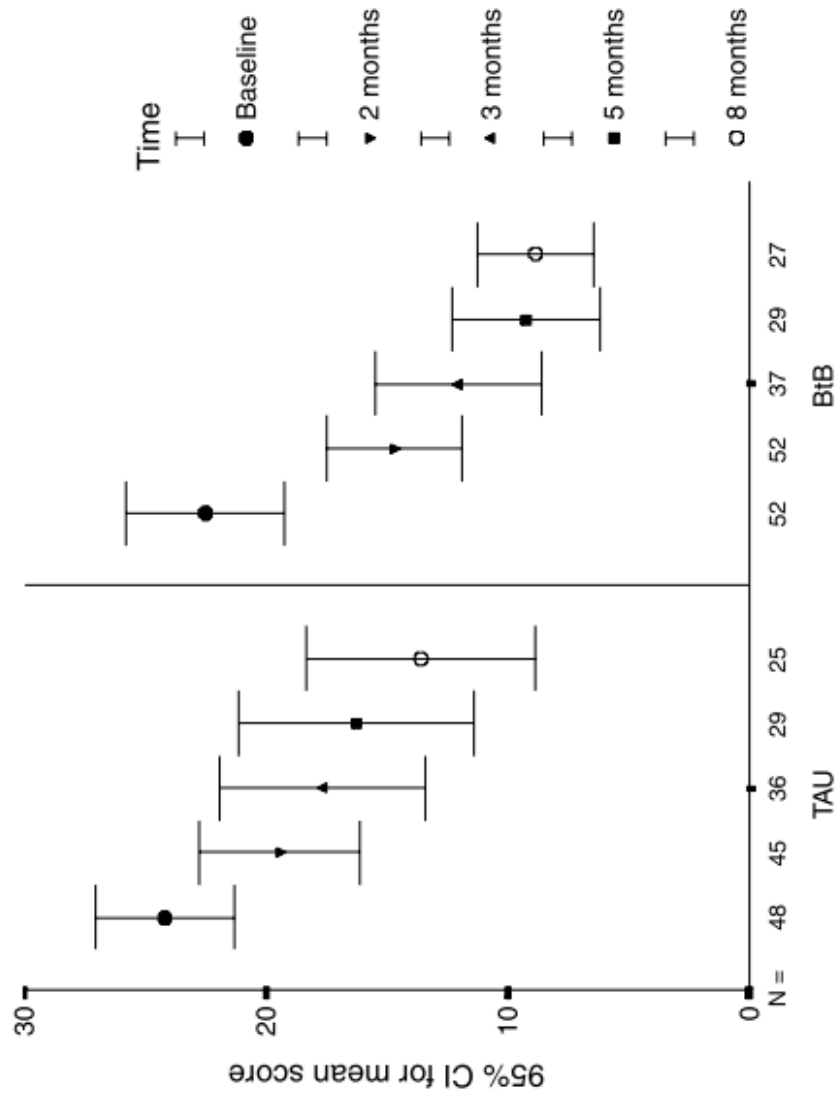
The data in Table 8.1 are repeated measures data with time as the single within-subject factor, i.e., they are longitudinal data. They have the following features that are fairly typical of those collected in many clinical trials in psychology and psychiatry:

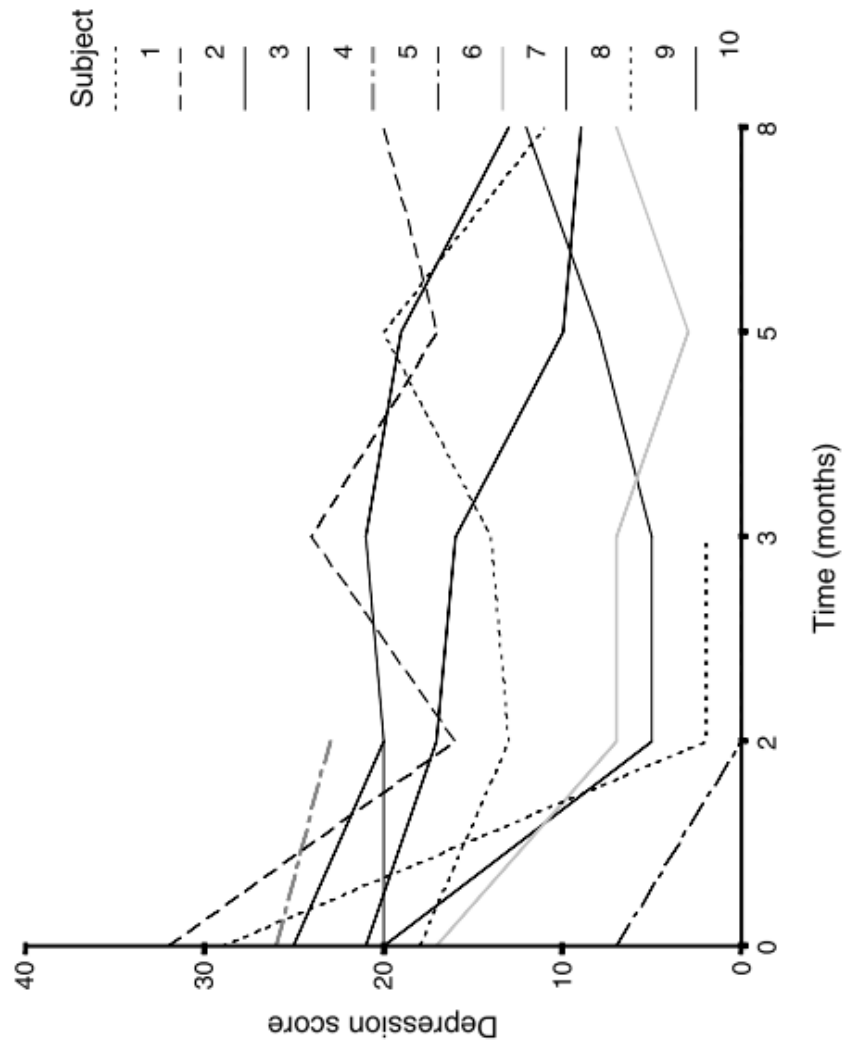
- There are a considerable number of missing values caused by patients dropping out of the study (coded 999 in Table 8.1).
- There are repeated measurements of the outcome taken on each patient post treatment, along with a baseline pre-treatment measurement.
- There is interest in the effect of more than treatment group on the response; in this case, the effect of duration of illness categorized as whether a patient had been ill for longer than six months (duration = 1) or for less than six months (duration = 0).

The question of most interest about these data is whether the BtB program does better than TAU in treating depression.

Table 8.1 A Subset of the Data from the Original BtB Trial

<i>Subject</i>	<i>Duration^a</i>	<i>Treatment^b</i>	<i>Depression Score</i>				
			<i>Baseline</i>	<i>2 Months</i>	<i>3 Months</i>	<i>5 Months</i>	<i>8 Months</i>
1	1	0	29	2	2	999 ^c	999
2	1	1	32	16	24	17	20
3	0	0	25	20	999	999	999
4	1	1	21	17	16	10	9
5	1	1	26	23	999	999	999
6	0	1	7	0	0	0	0
7	0	0	17	7	7	3	7
8	1	0	20	20	21	19	13
9	0	1	18	13	14	20	11
10	1	1	20	5	5	8	12
11	1	0	30	32	24	12	2





Mixed Model

→

Dependent Variable

bdipost

→

Factors

treat

time

duration

→

Covariates

bdipre

→

Cluster variables

subject

Estimation

Confidence Intervals

☒ REML

☒ Confidence intervals

Interval

95

%

Conditional on treatment and visit mean post-treatment, BDI is estimated to increase by 0.62 points for each additional baseline BDI score (95% CI from 0.46 to 0.77 points) and conditional on pre-treatment score and treatment group to decrease 0.7 points each month (CI from 0.41 to 0.99 points). Finally, for a given visit and baseline BDI score the BtB treatment is estimated to decrease the mean BDI score by 3.22 points (CI from 0.015 points increase to 6.5 points decrease).

Fixed Effects

Components

treat

time

duration

bdipre

Model Terms

bdipre

treat

time

duration

Fixed Intercept

✓

Random Effects

Components

treat : duration | subject

time : duration | subject

treat : bdipre | subject

time : bdipre | subject

duration : bdipre | subject

treat : time : duration | subject

treat : time : bdipre | subject

treat : duration : bdipre | subject

time : duration : bdipre | subject

treat : time : duration : bdipre | subj...

Random Coefficients

Intercept | subject

Effects correlation

Tests

Random intercept model

$$y_{ij} = \beta_0 + \beta_1 \text{Group}_i + \beta_2 t_j + u_i + \varepsilon_{ij}$$

Random intercept and slope model

$$y_{ij} = \beta_0 + \beta_1 \text{Group}_i + \beta_2 t_j + u_{i1} + u_{i2} t_j + \varepsilon_{ij}$$

Fixed Effect Omnibus tests

	F	Num df	Den df	p
bdipre	60.876	1	98.4	<.001
treat	3.893	1	93.9	0.051
time	7.968	3	191.3	<.001
duration	0.149	1	96.8	0.700

Note. Satterthwaite method for degrees of freedom

Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	14.787	0.8297	13.161	16.4133	98.4	17.823	< .001
bdipre	bdipre	0.610	0.0781	0.457	0.7628	98.4	7.802	< .001
treat1	1 - 0	-3.223	1.6334	-6.424	-0.0214	93.9	-1.973	0.051
time1	2 - 1	-1.439	0.8191	-3.045	0.1662	195.3	-1.757	0.080
time2	3 - 1	-2.671	0.8933	-4.422	-0.9200	197.0	-2.990	0.003
time3	4 - 1	-4.398	0.9291	-6.219	-2.5768	197.1	-4.733	< .001
duration1	1 - 0	0.646	1.6716	-2.630	3.9221	96.8	0.386	0.700

Random Components

Groups	Name	SD	Variance	ICC
subject	(Intercept)	7.25	52.6	0.674
Residual		5.05	25.5	

Note. Number of Obs: 280 , groups: subject 97

Conditional on treatment and visit mean post-treatment, BDI is estimated to increase by 0.61 points for each additional baseline BDI score (95% CI from 0.46 to 0.76 points) and conditional on pre-treatment score and treatment group to decrease 0.7 points each month (CI from 0.41 to 0.99 points). Finally, for a given visit and baseline BDI score the BtB treatment is estimated to decrease the mean BDI score by 3.22 points (CI from 0.015 points increase to 6.5 points decrease).

Conditional on treatment and visit mean post-treatment, BDI is estimated to increase by 0.62 points for each additional baseline BDI score (95% CI from 0.46 to 0.77 points) and conditional on pre-treatment score and treatment group to decrease 0.7 points each month (CI from 0.41 to 0.99 points). Finally, for a given visit and baseline BDI score the BtB treatment is estimated to decrease the mean BDI score by 3.22 points (CI from 0.015 points increase to 6.5 points decrease).